

Round 3 results — Run sparse-dense-0.45 submitted from CIR

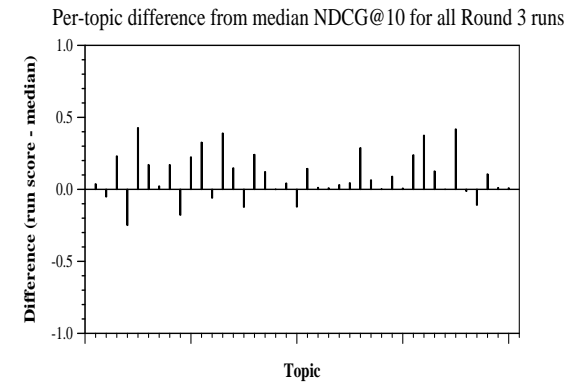
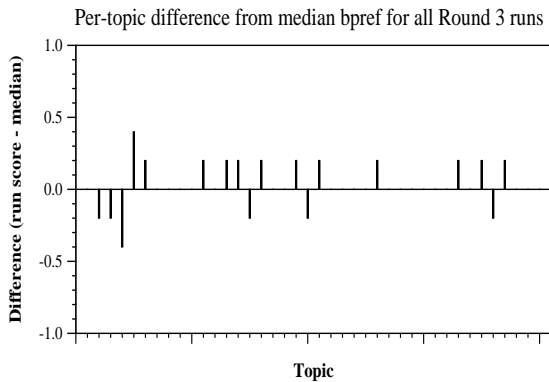
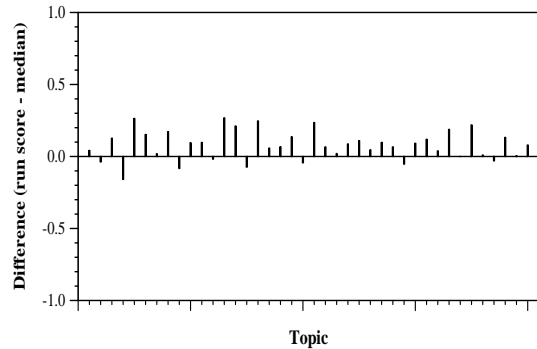
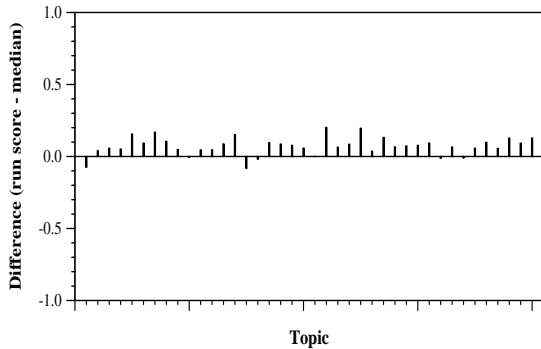
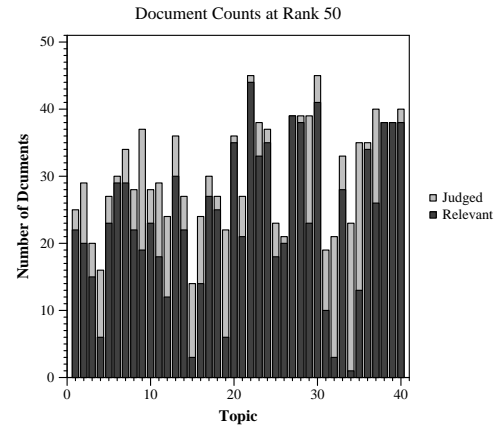
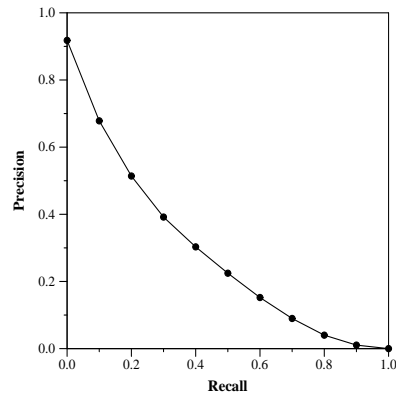
Run Description

Fusion of: (1) Anserini BM25 fusion baseline(RRF(2,4,6)), (2) ANN dense embedding retrieval trained on Med-MARCO (fusion weights tuned). Language model finetuned by performing MLM task on CORD19 Rnd3 documents.

Summary Statistics	
Run ID	sparse-dense-0.45
Topic type	feedback
Contributed to judgment sets?	no

Overall measures	
Number of topics	40
Total number retrieved	39923
Total relevant	4698
Total relevant retrieved	3109
MAP	0.2778
Mean Bpref	0.5364
Mean NDCG@10	0.6698
Mean RBP(p=0.5)	0.7400 +0.0539

Document Level Averages	
	Precision
At 5 docs	0.7450
At 10 docs	0.7050
At 15 docs	0.6650
At 20 docs	0.6200
At 30 docs	0.5708
R-Precision	
Exact	0.3272



Per-topic difference from median P@5 for all Round 3 runs

Per-topic difference from median RBP(p=0.5) for all Round 3 runs